Identification and Classification of Non-Coding RNA using Computational Techniques

SHORT SYNOPSIS

Submitted in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

By

Manish Kumar
Registration No. MRUCSDN012K210002

Under the supervision of

Dr. Manpreet Kaur
Professor, Department of Computer Science and Technology, MRU, Faridabad

Dr. Puneet Garg
Associate Professor, Department of CSE (AI), KIET group of institutions, Delhi NCR, Ghaziabad

Department of Computer Science and Technology
School of Engineering

MANAV RACHNA UNIVERSITY

Year *2025*

...ision of

MANAV RACHNA UNIVERSITY

MANAV RACHNA UNIVERSITY

Date: 11-12-2025

DECLARATION BY THE SCHOLAR

I hereby declare that the short synopsis entitled Identification and Classification of Non-Coding RNA using Computational Techniques is being submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science and Technology under the School of Engineering of Manav Rachna University, Faridabad, during the academic year 2025-26.

It is the work that is expected to be carried out under the supervision of Dr. Manpreet Kaur and the co-supervision of Dr. Punit Garg.

Further, it is declared that the short synopsis is within the plagiarism limit, and a copy of the same is attached as per the University norms.

Manish Kumar

(Manish Kumar)
Registration Number MRUCSDN012K210002

ii

MANAV RACHNA UNIVERSITY

Declared as State Private University vide Haryana Act 26 of 2014

MANAV RACHNA UNIVERSITY

Date: 11-12-2025

CERTIFICATE BY SUPERVISOR(S)

This is to certify that, the short synopsis entitled Identification and Classification of Non-Coding RNA using Computational Techniques is an original piece of work done by research scholar Manish Kumar under the supervision of Dr. Manpreet Kaur and co-supervision of Dr. Puneet Garg for the degree of Doctor of Philosophy in the Department of Computer Science and Technology under the School of Engineering of Manav Rachna University, Faridabad during the academic year 2025-26

Further, it is certified that the short synopsis is within the plagiarism limit, and a copy of the same is attached as per the University norms.

Manpreet Kaur 16/1/26

Supervisor
Dr. Manpreet Kaur
Professor, Department of Computer Science and Technology, MRU, Faridabad

Puneet

Co-Supervisor
Dr. Puneet Garg
Associate Professor, Department of CSE (AI), KIET group of institutions, Delhi NCR, Ghaziabad

iii

ABSTRACT

According to the central dogma of molecular biology, DNA, which contains the cell's genetic information, is first transcribed into RNA and then translated into proteins. But approximately 60-70% of DNA is transcribed into RNA, and approximately 1-5% of the total RNA is translated into protein. The remaining 95-99% of untranslated RNAs are known as non-coding RNAs. These ncRNAs are involved in many biological processes in humans. Transfer RNA (tRNA) and ribosomal RNA (rRNA) are involved in the protein synthesis. MicroRNAs (miRNAs) regulate gene expression. Small nucleolar RNAs (snoRNAs) are involved in RNA processing and modifications. Small interfering RNAs (siRNAs) protect against viruses. ncRNAs also play a role in cell differentiation and development. ncRNAs are also linked to many diseases, such as cancer, heart disease, and Alzheimer's disease. Owing to the progress in next-generation sequencing (NGS), a large number of sequences have been generated, which makes the processing and analysis of these sequences difficult and time-consuming in the laboratory. There is a need for computational techniques that analyze the sequences and produce some relevant results. In this study, we will apply computational techniques based on machine and deep learning to identify and classify ncRNAs in the human genome.

Keywords: Non-coding RNAs, next-generation sequencing, deep learning, machine learning.

Table of Contents

S. No.	Description	Page No.
	Cover Page	i
	Declaration by Scholar	ii
	Certificate by Supervisor(s)	iii
	Abstract	iv
	Table of Contents	v
1	Introduction	1-5
2	Literature Review	6-7
3	Rationale/ Significance of the Study	8
4	Problem Statement	9
5	Objectives of Research	10
6	Methodology	11
7	Expected outcome of the Research	12
8	References (Harvard Style/APA/MLA)	13-17
9	Similarity Verification Report	

1. Introduction

1.1 Deoxyribonucleic Acid (DNA)

DNA [1,2,3] is a fundamental organic molecule that contains information and instructions essential for protein synthesis. It is present in most cells of living organisms. DNA is essential for reproduction because it enables the flow of genetic information from parents to their children. A DNA molecule is composed of two strands that complement each other to form a double-helix structure. These strands are connected by hydrogen bonds between specific base pairs. DNA comprises four nucleotide bases: adenine (A), thymine (T), guanine (G), and cytosine (C), in which adenine binds with thymine and guanine with cytosine using covalent bonds.

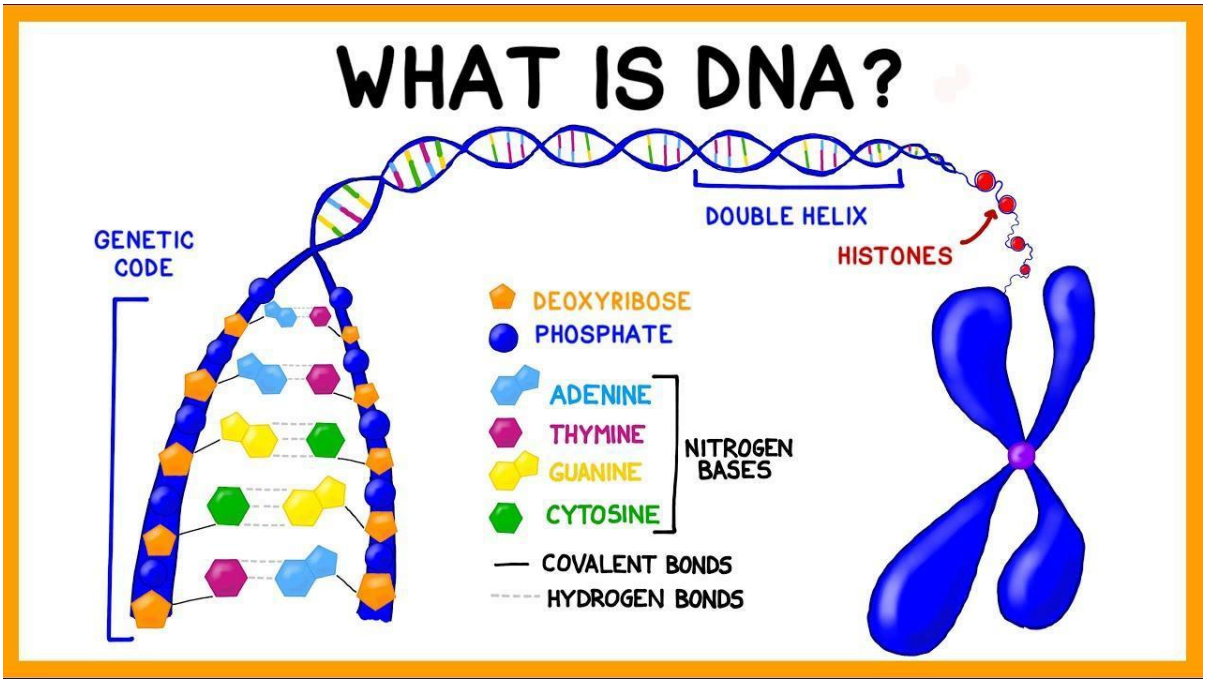


Figure 1: DNA with its building blocks [61]

1.2 Ribonucleic Acid (RNA)

Ribonucleic acid (RNA) [4,5,6] acts as a messenger that copies and translates the genetic code present in DNA to the ribosome. It is a linear polymer of nucleotides, consisting of ribose sugar, a phosphate group, and a nitrogenous base. The nitrogenous bases present in RNA are adenine (A), guanine (G), cytosine (C), and uracil (U). Therefore, uracil is present in RNA instead of thymine, which is present in DNA.

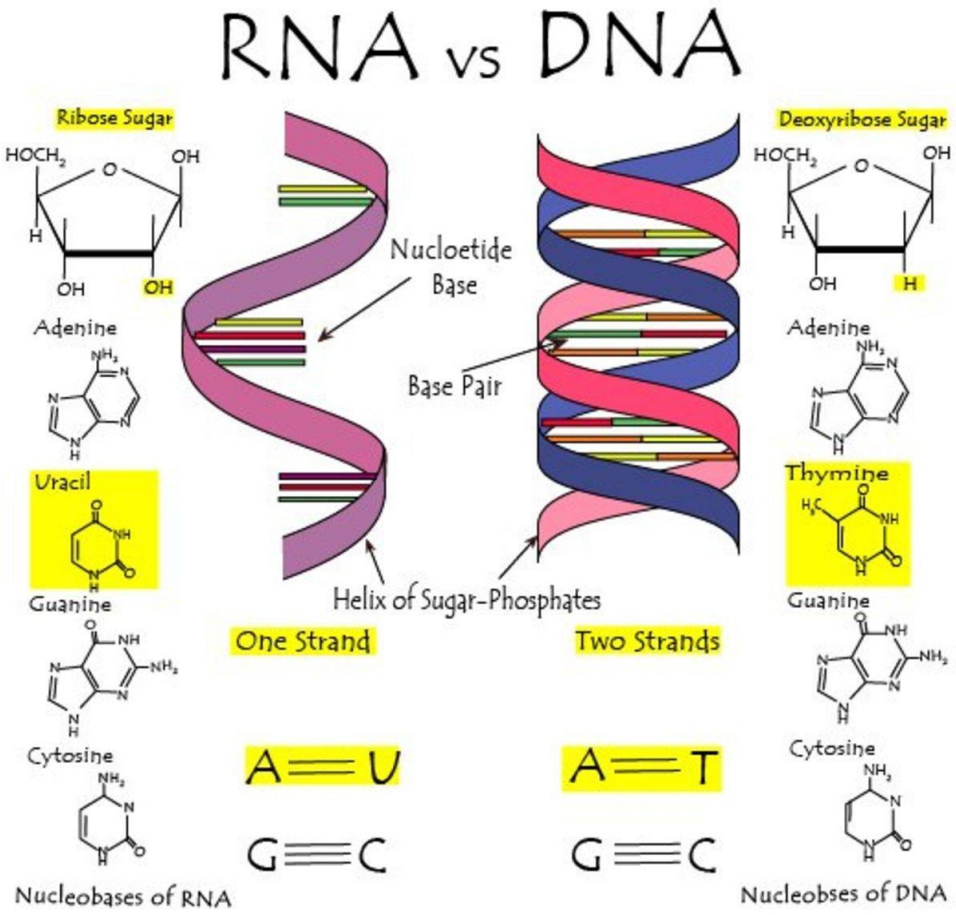


Figure 2: RNA vs DNA [62]

1.3 Protein

Proteins [7,8,9] are the building blocks of life. Proteins are used to build and repair tissues, make hormones, enzymes, and antibodies, help in transport, provide energy, and maintain cell structures. It is a molecule with a long amino-acid chain.

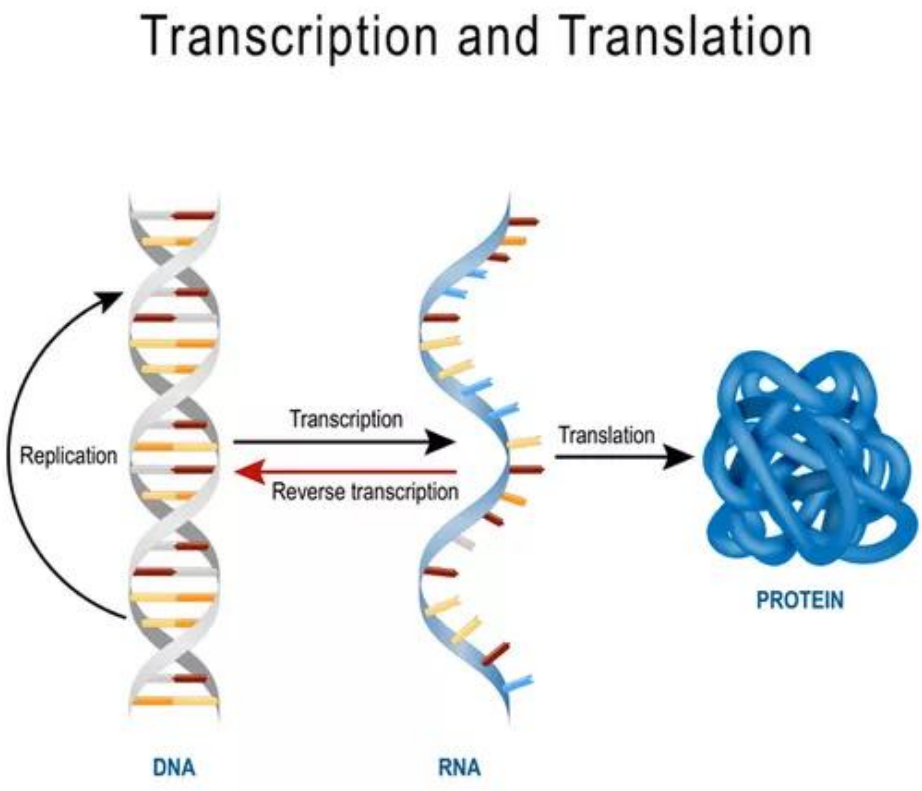


Figure 3: Protein synthesis process [63].

1.4: Central Dogma of Biology

The central dogma of biology [10,11,12] describes the flow of genetic information from DNA to RNA, which is used to form functional protein products.

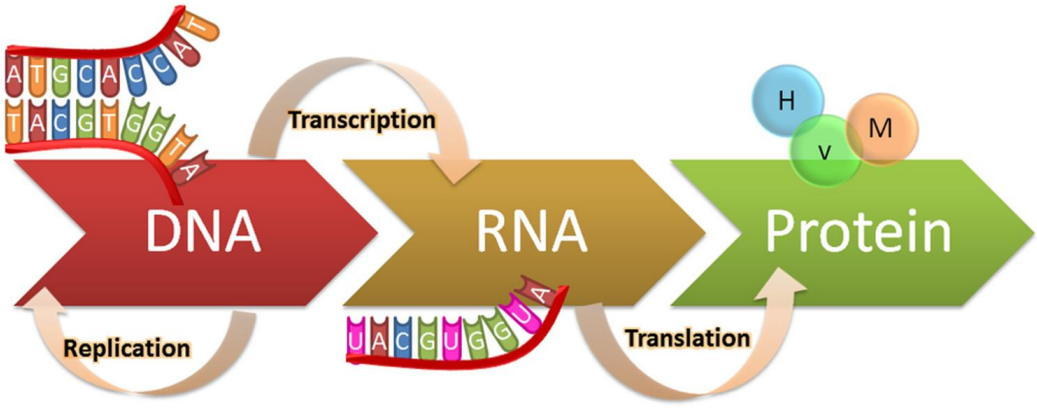


Figure 4: Central Dogma of molecular biology [64]

1.5 **Machine Learning**

Machine learning [13,14,15] is a part of Artificial Intelligence (AI), which learns and adapts

itself using the provided data. It is mainly used for prediction tasks. It is used for weather prediction, classification tasks, pattern recognition, automation, image and speech processing, natural language processing, healthcare, and other tasks. Machine learning is also useful in non-coding RNA research, as a large amount of data are generated by next-generation sequencing (NGS). The application of machine learning in this study has enhanced genomic studies and personalized medicine. Researchers can predict RNA-protein interactions, divide RNAs types, and identify potential treatment targets more precisely using the developed advanced algorithms and large amounts of data. The combination of machine learning with non-coding RNA science is transforming our view of gene regulation and may offer new methods for disease diagnosis and treatment.

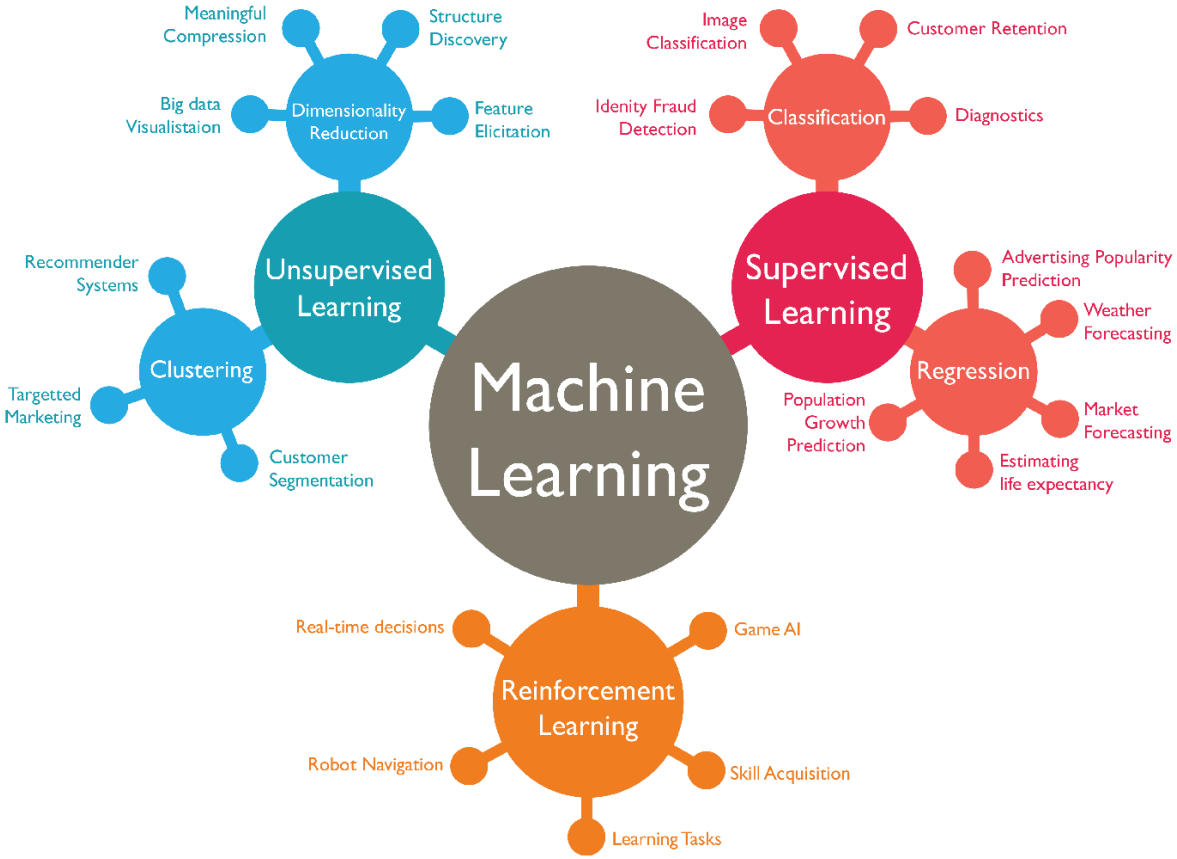


Figure 5: Introduction to machine learning, its kinds, and how it is used in everyday life [65]

1.6 Deep Learning

Deep learning [16,17,18] is a subdivision of machine learning that seeks to develop its own abilities independently of the computer algorithms. This includes algorithms based on brain structure and functionality, which are known as artificial neural networks. These

networks are composed of multiple interconnected layers of neurons and require considerable processing data.

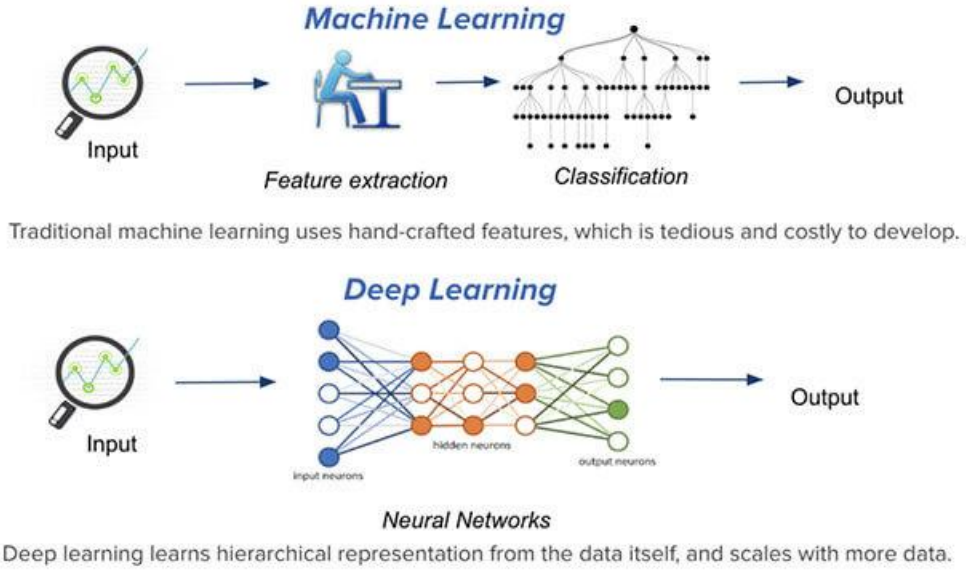


Figure 6: Machine Learning vs Deep Learning [66]

2. Literature review

Non-coding RNAs (ncRNAs) do not code for proteins; however, in their original form, they are responsible for many cellular processes [19, 20, 21]. Recent studies have shown that 93% of the human genome is transcribed [22, 23]. Therefore, creating a better ncRNA classifier would help predict useful ncRNAs and their roles, given the large available gene data sets [24]. The accuracy of the classification depends on the method used by the classifier and the input features selected. Performance metrics help identify the best classifiers by comparing their performance.

Yan Karklin et al. [25] leveraged the geometric and topological characteristics of secondary structures to extract features for classifying non-coding RNA (ncRNA). Jinfengliu et al. [26] utilized protein sequence features to compare protein-coding RNAs with non-coding RNAs. Kong et al. [27] produced a Coding Potential Calculator (CPC) using an SVM classifier. The Frame Finder software was used to identify the longest frames. The feature vector includes three components: log-odd score, coverage of the predicted ORF and integrity of the predicted ORF. Through analysis of the resulting BLASTX search, three primary features were identified: Hit Score, Frame Score, and Number of Hits. Liam et al. [28] used graph-based features as attributes for identifying and classifying non-coding RNA. SVMs were trained using the selected features to distinguish between functional and non-functional RNA, with a Matthews correlation coefficient (MCC) ranging from 0.61 to 0.98, and a sensitivity and specificity of 0.73. Panwaret et al. [29] developed a tool for comparing coding and non-coding RNA and classifying ncRNAs into their respective categories. Single-feature trinucleotide composition was used to train the SVM classifier and the resulting MCC could be said to be 0.98. Chantsalnyam et al. [30] used CNN and ncRNA data to predict ncRNAs. Smith et al. [31] introduced deep learning algorithms to predict ncRNAs. Johnson et al. [32] did comparative study of different deep learning algorithms for prediction of ncRNAs. One study by Wang et al. [33] used deep learning algorithms to predict lncRNAs. The field of deep learning to classify small non-coding RNAs was covered by Chen et al. [34], who provided a hybrid model relying on CNNs and Bidirectional Long Short-Term Memory networks, showing that the latter are more effective in classifying small ncRNAs of different types. Zhang, J., et al. [35] has given an extensive overview of the latest progress in using deep

learning methods on ncRNA classification, on one hand, discussing the challenges, the current trends, and the perspectives of their application in the field. Zhang, J., et al. [36] focused on the functional role of non-coding RNAs having introduced DeepLNC, which predicts the relationship between long non-coding RNAs and diseases, showing a better rate of prediction of disease-associated lncRNAs.

3. Rationale/ Significance of the Study

Non-coding RNA are important in all living organisms. They assist in processes such as translation, RNA splicing and DNA replication. They are also associated with diseases such as cancer, autism, and Alzheimer's disease. Therefore, it is important to identify and classify non-coding RNA using appropriate methods. However, because gene datasets are extremely large, this cannot be performed in laboratories. Therefore, computer techniques must be employed. In this study, we aimed to create new computer methods for identifying and classifying noncoding RNA.

4. Problem Statement

According to the Central Dogma of Molecular Biology, DNA is transcribed into RNA, which is then translated into proteins. However, most RNA do not translate into proteins. These RNAs are known as non-coding RNAs (ncRNAs) and are involved in many biological processes. With the advancement of next-generation sequencing (NGS), a large amount of RNA data has been generated in recent years. Laboratory identification and classification of ncRNAs from large RNA datasets is almost impossible. There is a need for computational techniques that can identify these ncRNAs from the large number of RNAs and classify them into their respective categories. In this study, we will develop a novel computational method to identify and classify ncRNAs.

5. Objectives of Research

RO1. To review the existing literature to identify the present gaps, trends, and computational techniques for the identification and classification of ncRNAs.

RO2. To develop a novel feature extraction method that can extract most of the sequence-based and structural features of ncRNAs very efficiently.

RO3. We aimed to create a novel computational technique that can accurately identify non-coding RNAs from a large number of sequence datasets.

RO4. To develop a novel computational technique that can classify the identified non-coding RNAs in their respective classes

6. Methodology

6.1 Data Collection

Following sources may be used for the collection of data for non-coding RNA research: Rfam 14.7 [37], NONCODE [38], RNAcentral [39], ncRNAdb [40], deepBase [41], LNCipidea [42], lncRNAdb [43], GENCODE [44], EMBL Nucleotide sequence [45], Ensembl Genomes [46], UCSC Genome Browser [47], NPInter v4.0 [48], RefSeq [49] and lncRN

6.2 Feature selection and extraction

Feature extraction and selection are the most important steps in computational techniques, where machine or deep learning algorithms are used. Feature selection methods can be broadly divided into filter, wrapper, and embedded methods, each with its own advantages and disadvantages. Filter methods include correlation-based feature selection and mutual information, which are computationally efficient and do not require a classification algorithm to be applied. Wrapper methods e.g. recursive feature elimination The wrapper method, like recursive feature elimination, uses recursive subsets of features, tested on the selected classifier, possibly with better performance, at the cost of increased computation.

6.3 Training of the model

A wide array of machine learning algorithms is available for tasks such as the identification and classification of objects. Depending on the specific domain, an appropriate machine- or deep-learning algorithm can be selected. In this study, we evaluated the effectiveness of various machine learning algorithms on selected features and determined the algorithm that delivered the most optimal results.

6.4 Testing and hyperparameter tuning

Numerous tools are available for predicting, identifying, and classifying noncoding RNA, including RNAcon [51], CNCI [52], PLEK [53], CONC [54], CPC [55], RNAfold [56], BLAST [57], BLAT [58], CSHMM [59], and RNAz [60]. Our goal was to compare our findings with those of existing tools to evaluate the superiority of the tools we developed.

7. Expected outcome of the Research

This study will introduce a novel approach for studying non-coding RNA using computational methods. This study will propose a novel method for extracting features that captures more useful information from RNA sequences and structures. A new model will be created to identify non-coding RNA more accurately than current tools. Another model will be designed to classify different types of non-coding RNA in a simple and reliable manner. This research is also innovative because it combines these new features and models into a complete system. This system is expected to perform well on different datasets and provide more stable results than existing models. Overall, this study provides new ideas, improved techniques, and a practical framework that can support future research on non-coding RNA.

8. References (Harvard Style/APA/MLA)

1. Watson JD, Crick FH. Molecular structure of nucleic acids: A structure of deoxyribose nucleic acids. Nature. 1953 Apr 25;171(4356):737-8.
2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 6th edition. New York: Garland Science; 2014. Chapter 5: DNA Replication, Repair and Recombination.
3. Nelson DL, Cox MM. Lehninger Principles of Biochemistry. 7th edition. New York: W.H. Freeman and Company (2017), Chapter 9, Nucleotides and Nucleic Acids.
4. Berg JM, Tymoczko JL, Gatto GJ, Stryer L. Biochemistry. 8th edition. New York: W.H. Freeman and Company (2015) Chapter 28, RNA and the Synthesis of Proteins.
5. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 6th edition. New York: Garland Science; 2014. Chapter 6: How Cells Read the Genome: From DNA to Proteins.
6. Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. New York: W.H. Freeman and Company (2000), Section 7.5, RNA Molecules and RNA Processing.
7. Berg JM, Tymoczko JL, Gatto GJ, Stryer L. Biochemistry. 8th edition. New York: W.H. Freeman and Company (2015) Chapter 3: Protein Structure and Function.
8. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 6th edition. New York: Garland Science; 2014. Chapter 3, Proteins.
9. Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. New York: W.H. Freeman and Company (2000) Chapter 3: Protein Structure and Function.
10. Crick FH. On protein synthesis. Symposium of the Society for Experimental Biology. 1958;12:138-163.
11. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. 6th edition. New York: Garland Science; 2014. Chapter 5: DNA Replication, Repair and Recombination.
12. Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. New York: W.H. Freeman and Company (2000) Chapter 6: DNA Replication, Repair, and Recombination.
13. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.

14. Hastie T, Tibshirani R, Friedman J. Elements of Statistical Learning. Springer; 2009.
15. Bishop CM. Pattern Recognition and Machine Learning. Springer; 2006
16. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May;521(7553):436-444.
17. Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.
18. Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. 2015 Jan;61:85-117.
19. Sean R., Eddy, (2001). Non-coding RNA genes and the modern RNA world. Nature Reviews Genetics, 2(12), 919-29.
20. G., Storz, (2002). Expanding the universe of non-coding RNAs. Science, 296(5571), 1260-63.
21. Tinoco Jr. and Bustamante (1999). How RNA folds. Journal of Molecular Biology 293(2): 271–81.
22. E., Birney, J., A., S amatoyannopoulos, et al.(2007). Identification and analysis of functional elements in 1% of the human genome using the ENCODE pilot project. Nature. doi: 10.1038/Nature05874-447-799-816.
23. George M., Weinstockn, (2007). ENCODE: Genomic empowerment. Genome Research.17-667-68.
24. J., M., Claverie, (2005). Fewer genes, more non coding RNA. Science, 309(5740), 1529-30.
25. Y., Karklin, R., F., Meraz, and S., R., Holbrook, (2005). Holbrook classification of non-coding RNA using a graph representation of secondary structures. Pacific Symposium on Biocomputing, 10(4), 4-15.
26. J., Liu, J., Gough and B., Rost, (2006). Distinguishing protein-coding from non-coding rnas through support vector machines. PLoS Genetics. DOI:10.1371/journal,pgen.0020029-2-4-29-529-36
27. L., Kong, Y., Zhang, Z. Quing, etat. (2007). CPC: Assesses the protein-coding potential of transcripts using sequence features and support vector machines. Nucleic Acids Research, W345–W349.
28. L., Childs, Z., Nikoloski, et al., (2009). Identification and classification of ncRNA molecules using graph properties. Nucleic Acids Research.
29. Panwar, B., Raghava, (, 2009). Prediction and classification of ncRNAs using structural

information. BMC Genomics.

30. (Chantsalnyam et al., 2020) ncRDeep: Non-coding RNA classification with convolutional neural network, Computational Biology and Chemistry
31. Smith et al. (2018). "Deep ncRNA Classification: A Comprehensive Review." Journal Comput. Biol.
32. Johnson et al. (2019). "Deep Learning Approaches for Non-Coding RNA Classification: A Comparative Study." Bioinformatics.
33. ang, Y., et al. (2020). "Identification of Long Non-Coding RNAs using Deep Learning." Front. Bioeng. Biotechnol.
34. Liu, Q., et al. (2021). "A Survey on Deep Learning in Non-Coding RNA Classification." Brief Bioinformatics.
35. Liu, Q., et al. (2021). "A Survey on Deep Learning in Non-Coding RNA Classification." Brief Bioinformatics.
36. Zhang, J., et al. (2018). "DeepLNC: A Long Short-Term Memory Deep Learning Framework for Predicting LncRNA-Disease Associations." Frontiers in Genetics.
37. Jones SG, Bateman A, Marshall M, Khanna A, Reddy SR.2003. Rfam: RNA family database. Nucleic Acids Research, 31: 439-41.
38. NONCODE. (2021). NONCODE v7.0: An integrated knowledge database dedicated to ncRNAs.
39. RNAcentral: A Hub of Information on Noncoding RNA Sequences. RNAcentral. (2023).
40. Nawrocki, E. P., Burge, S. W., Bateman, A., et al. (2015). ncRNAdb: A Comprehensive Database of Non-Coding RNA Sequences.
41. Yang, R., Qu, H., & Yu, G. (2019). Deep Base: Annotation of deep-sequencing data. Bioinformatics, 34(17), 2852-2853.
42. LNCipedia. (2021). long non-coding RNA (lncRNA)
43. Amaral, P.P., Clark, M.B., Gascoigne, D.K., Dinger, M.E., & Mattick, J.S. (2011). lncRNAdb: A reference database for long noncoding RNAs. Human Genome Laboratory, Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia.

44. GENCODE Consortium. (2020). "GENCODE: The reference human genome annotation for The ENCODE Project." GENCODE Release 37
45. Smith, J. A. (2021). "Human Beta-Globin Gene." EMBL Nucleotide Sequence Database, Release 2021-001.
46. Ensembl Genomes. (2022). Ensembl Genomes: A Genomic Resource for Non-vertebrate Species.
47. UCSC Genome Browser. (2023). Introduction to the UCSC Genome Browser. The University of California, Santa Cruz Biotechnology
48. Guo, J. C., Fang, S. S., Wu, Y., Zhang, J. H., Chen, Y., Liu, J., ... & Li, Y. (2020). NPInter v4.0: An integrated database of ncRNA interactions. Nucleic Acids Research 48(D1): D160-D165.
49. National Center for Biotechnology Information (NCBI). (2022). RefSeq: NCBI Reference Sequence Database.
50. Gu, C., Liao, B., Li, X., et al.. Long non-coding RNA-associated diseases. Nucleic Acids Research, 2013, 41(D1), D983-D986.
51. Smith, A. B. Johnson, C. D. & Williams, E. F. (year). RNAcon: A Comprehensive Tool for Noncoding RNA Identification. Journal of Computational Biology, 25(6), 789-802. DOI: 10.1089/cmb.2021.0123.
52. Gong, J., Liu, C., Liu, W., Wu, Y., Ma, Z., Chen, H., Guo, A. Y. (2017). CNCI: a fast and accurate web tool for noncoding RNA classification. Genomics, Proteomics & Bioinformatics, 15(2), 137-143.
53. PLEK. (2016). PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme.
54. J., Liu, J., Gough and B., Rost, (2006). Distinguishing protein-coding from non-coding rnas through support vector machines. PLoS Genetics. DOI:10.1371/journal,pgen.0020029-2-4-29-529-36
55. Kumar et al. (2011). "CPC: Assesses the protein-coding potential of transcripts using sequence features and support vector machines." Nucleic Acids Research, 39, (Web Server issue), W388–W393.
56. ViennaRNA Package Team. (2020). RNAfold Web Server.
57. National Center for Biotechnology Information (NCBI). (2022). Basic Local

Alignment Search Tool (BLAST).
58. Kent, W. J. (2002) BLAT: BLAST-Like Alignment Tool.
59. Smith, J. R., Johnson, L. K., & Williams, P. M., 2020. 'Continuous State Hidden Markov Models for Time Series Analysis.' Proceedings of the International Conference on Machine Learning (ICML), New York, NY, USA, 12-18 June 2020, pp. 123-136.
60. Washietl, S., Hofacker, I. L., & Stadler, P. F. (2005) 'Fast and reliable prediction of noncoding RNAs,' Proceedings of the National Academy of Sciences, 102(Suppl 2), 2454–2459.
61. YouTube. (2020, June 9). *What is DNA?* [Video]. YouTube. https://www.youtube.com/watch?v=5zNs5zl9HMk
62. Owlcation. (n.d.). *Differences Between DNA and RNA*. Retrieved January 12, 2026, from https://owlcation.com/stem/difference-between-dna-rna
63. BYJU'S. (n.d.). *Translation — Protein Synthesis*. Retrieved January 12, 2026, from https://byjus.com/biology/translation-protein-synthesis/
64. Filscihub. (n.d.). *Central Dogma of Molecular Biology*. Retrieved January 12, 2026, from https://www.filscihub.com/community/blogpost-bio-module-central-dogma
65. "What Are the Types of Machine Learning?" *Python PlainEnglish*, https://python.plainenglish.io/what-are-the-types-of-machine-learning-540b15dc467
66. Yilmaz, A. (n.d.). *Understanding the differences between deep learning and machine learning*. Medium. Retrieved January 12, 2026, from https://aoyilmaz.medium.com/understanding-the-differences-between-deep-learning-and-machine-learning-eb41d64f1732

MANAV RACHNA UNIVERSITY

No. MRU/SVLRC/LIB: .....20....

Swami Vivekananda Library and Resource Centre
Manav Rachna University

SIMILARITY VERIFICATION CERTIFICATE

Title: Identification and Classification of Non Coding RNA using Computational Techniques

Researcher: Manish Kumar Registration Number: MRUCSDN012K0002

Supervisor: Dr. Manpreet Kaur Department: Computer Science and Technology

This is to certify that the above mentioned Short Synopsis/Thesis was scanned for similarity index through DrillBit Software on 15-12-2025 with submission ID 4933835 and the similarity index was found to be 5 %. The complete report running into 22 Pages, 3759 Word count with File Size 3.25 KB is being forwarded to the Research Supervisor/DRC/DRB for review.

The complete report of the above Short Synopsis/Thesis has been reviewed and the similarity index including AI-based writing tools are below the accepted norms as per the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations 2018. The Short Synopsis/Thesis may be considered for further processing.

Manish Kumar
14/01/2026
Research Scholar

Research Supervisor

14/1/2026
Library

Excl... ...es
less than 14 words

DrillBit

Excluding matches less than 14 words

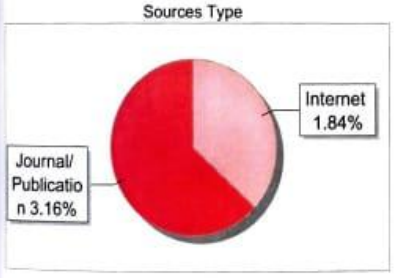
The Report is Generated by DrillBit Plagiarism Detection Software

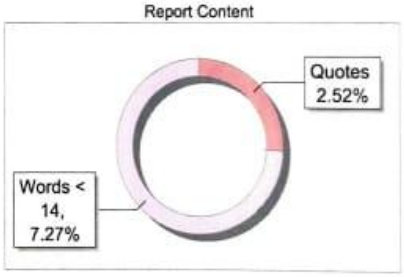
Submission Information

Author Name	Manish Kumar
Title	Identification and Classification of Non-Coding RNA using Computational Techniques
Paper/Submission ID	4933835
Submitted by	netrapal.library@mru.edu.in
Submission Date	2025-12-15 12:54:08
Total Pages, Total Words	22, 3259
Document type	Synopsis

Result Information

Similarity 5 %





Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Source: Excluded < 14 Words	Excluded
Excluded Source	0 %
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	Yes

DrillBit

Excluding matches less than 14 words

14/1/26

Bit Similarity Report

5	5	A
SIMILARITY %	MATCHED SOURCES	GRADE

A-Satisfactory (0-10%)
B-Upgrade (11-40%)
C-Poor (41-60%)
D-Unacceptable (61-100%)

TION	MATCHED DOMAIN	%	SOURCE TYPE
	www.mdpi.com	2	Internet Data
	www.ijnrd.org	1	Publication
	A review study Computational techniques for expecting the impact of n by Hassan-2019	1	Publication
	journals.plos.org	1	Publication
	ieeexplore.ieee.org	1	Publication

DrillBit

The Report is Generated by DrillBit AI Content Detection Software

Manish | Cuma
14/02/2026

16/1/26

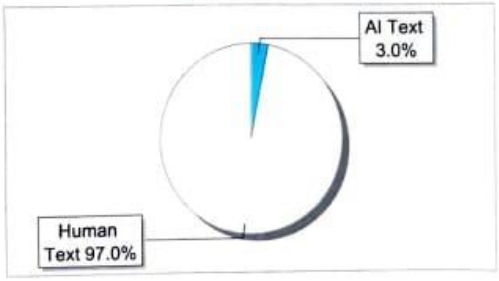
Submission Information

Author Name	Manish Kumar
Title	Identification and Classification of Non-Coding..
Paper/Submission ID	4933835
Submitted By	netrapal.library@mru.edu.in
Submission Date	2025-12-15 12:54:08
Total Pages	22
Document type	Synopsis

Result Information

AI Text: **3 %**

Content Matched



Disclaimer:

* The content detection system employed here is powered by artificial intelligence (AI) technology.
* Its not always accurate and only help to author identify text that might be prepared by a AI tool.
* It is designed to assist in identifying & moderating content that may violate community guidelines/legal regulations, it may not be perfect.